A Work Project, presented as part of the requirements for the Award of a Master's degree in Finance from the Nova School of Business and Economics.

Bankruptcy prediction using machine learning models: empirical results in the Colombian manufacturing industry (2018-2022)

Miguel Angel Parra Castro

Work project carried out under the supervision of:

Yeny Rodríguez

Virginia Gianinazzi

19/12/2023

Abstract

The purpose of this study is to examine financial indicators that reveal the situation of corporate failure in manufacturing companies in Colombia. The use of these indicators is based on previous studies that have used predictive models of corporate fragility: multiple discriminant analysis, logistic regression, and machine learning. This work uses logistic regression and random forests models. This work is based on financial indicators made of the data reported between 2018 and 2022 in the database Sistema de Información y Riesgos Empresariales (SIREM) of the Superintendence of Companies.

Keywords: Corporate finance, business failure, bankruptcy prediction, insolvency, financial analysis, financial ratios, random forest.

Introduction

A company is an economic entity that engages in commercial, industrial or service-providing activities, with the objective of generating economic profit (Sepúlveda, 2021). Gelashvili & Segovia-Vargas (2020) state that the function of a company is to generate value for its shareholders, maintain a profitable and sustainable operation, effectively manage its financial resources, and meet its financial and commercial obligations. The survival of firms contributes to economic growth, employment promotion, market competition and financial stability (Alarcón & Mejía, 2017; Umaña, 2012).

To ensure the survival of firms, financial analysis serves as a primary tool. Financial analysis of a firm comprises a thorough review to financial statements, such as balance sheet, profit and loss statement, cash flow statement, to monitor financial health, performance, and financial position in the industry (Faxas & Fuentes, 2011; Zeng, 2013). Thus, financial analysis includes the use of various financial ratios in order to measure aspects of liquidity, solvency, profitability and efficiency (Lakshmi et al., 2021; Rahim, 2021). In addition, financial management decisions such as capital structure, investment strategies and risk management are an essential factor to ensure the success and financial viability of a company in the long term (Faxas & Fuentes, 2011).

Finding the factors that give rise to corporate failure is one of the objectives of governments, guilds, financial agents, and society (Parra & Peluffo, 2022). Corporate failure generates a notable concern to all stakeholders of the company, competitors, investors, debtors, creditors, customers, employees, suppliers, and partners (Galán-Barrera & Torres-García, 2017). For that reason, this problematic has been the subject of various research and models that serve as predictors of situations that can be rectified.

In the field of finance, predicting bankruptcy has been considered as a relevant topic to investigate (Zięba et al., 2016). Research on financial stress based on financial indicators and statistical methods began in 1932 with Fitzpatrick who established the descriptive stage and then in 1968 Altman gave rise to the predictive stage with the formulation of multivariate models (Altman, 1968). Discriminant analysis and financial ratio analysis have been the main techniques used in several studies that have been carried out to measure their effectiveness in predicting corporate bankruptcy (Casanova, 2011; Correa & Mejia, 2019). Thus, Altman (1968) proposed the scope of financial ratio analysis as a statistical tool to anticipate a business crisis. In this way, Ohlson (1980) revealed empirical results in the prediction of business failure using the maximum likelihood estimation methodology of the conditional logit model.

Another research indicates that risk models based on survival analysis are more suitable for forecasting corporate fragility than models that only consider a single period of analysis, and a model that takes accounting indicators and market variables (market size, stock returns) as inputs to generate more accurate forecasts is configured (Shumway, 2001). By incorporating time-varying independent variables and considering the probability of bankruptcy occurring given a time window (Xu, 2019), these survival models are appropriate in the complex economic landscape where fragility risk is time-varying (Duda & Asgharian, 2010).

More recent studies use novel methodologies to anticipate business failure based on machine learning and artificial intelligence such as: neural networks (Atiya, 2001; Wilson & Sharda, 1994), random forests (Huang et al., 2017; Joshi et al., 2018; Sinelnikova-Muryleva et al., 2018; Garcia et al., 2019; Johnpaul et al., 2019; Viswanathan et al., 2020; Alam, et al., 2021; Gurnani, et al., 2021; Yousaf, et al.,2022; Sermpinis et al., 2023); support vector machines (Lin et al., 2011; Shin et al., 2005) and ensemble methods (Kim & Kang, 2010; Wang & Ma, 2012). In general, the latter techniques have better classification performance than conventional statistical models.

In this way, it results interesting analyze what financial indicators predict insolvency risk in the Colombian manufacturing sector using machine learning models? Therefore, the objective of the research is to estimate the probability of bankruptcy of companies in the manufacturing sector in Colombia for the period 2018-2022, based on financial ratios and using machine learning classification methodologies such as logistic regression and random forests.

Finally, the general structure of this work presents 4 parts. The first presents a comprehensive literature review about the topic. The second describes the methodologies used which will validate the analysis of the results of the estimation of financial insolvency in the companies of the manufacturing sector in Colombia for the period 2018-2022. The third contains the results of the estimations and their analysis. Finally, the conclusions of the work will be presented.

1. Literature Review

This section presents the definition of insolvency, then it reviews the financial indicators that have been identified as the main predictors of insolvency, and finally it presents the main methodologies that have been used in the estimation of state of bankruptcy.

- Definition of insolvency

The term business bankruptcy is a concept that can be interpreted in different ways, depending on the author or the legal perspective adopted. According to Gitman et al. (2015), the bankruptcy or failure of a company is an unfortunate situation that can occur for a variety of reasons, either due to poor management in the first years of operation or due to inadequate practices in larger companies (Bernate and Gómez, 2021).

Lev (1978) argues that business failure implies the generation of losses or bad practices in terms of financial profitability; Zacharakis et al. (1999) considers that a company has failed when it is forced to liquidate to avoid greater losses, as well as when the value of the company is lower than its operating cost. This definition of business failure focuses on the company's inability to generate profits and stay afloat in the market; in agreement Romero (2013), understands it as the inability to pay debts (Bernate and Gómez 2021). The conjuncture of the company over time is the basis for establishing the definition of insolvency (Galan-Barrera & Torres-Garcia 2017). The former when there is a temporary situation in which the company truthfully demonstrates that, due to market or internal factors of the company, serious difficulties could arise that prevent the fulfillment of its current obligations (López, 2015).

Some studies in the literature have identified financial indicators that determine the insolvency of firms which are analyzed below.

- Predicting bankruptcy through financial indicators

Financial statements are essential for decision making in business management because they provide detailed information about the financial situation of the firm (Prats et al., 2022) and allow evaluating its performance in terms of profitability and growth (Araica & Meda, 2017), making it easier to access financing and investment (Garay et al., 2019). The information extracted from financial statements serves as input to establish probabilistic models that support in the prediction of corporate bankruptcy (Garcia & Flores, 2010).

According to Andrade (2017), a financial ratio is obtained by comparing two accounts in the financial statements, which will provide us with exactly the relative information of one account with respect to another; depending on the measurement objective, the ratio can be understood in monetary units or expressed in percentage terms. The importance of financial indicators in avoiding business failure lies in their ability to provide decisive information on the financial health of a firm (García-Rosales & Chávez-Chávez, 2023). For example, indicators such as return on equity (ROE) and return on investment (ROI) allow assessing the efficiency with which the corporation employs its resources to generate profit, which in turn helps to recognize feasible financial complications and adopt well-founded actions to prevent bankruptcy (Talavera et al., 2016). In addition, the use of financial ratios in decision making, oriented in the study of the company's financial performance, provides results that contribute to management in making fundamental decisions for the sustainability and expansion of the business (Talavera et al., 2016).

Altman (1968) states that financial ratios have their origin in the relationship between variables in financial statements and a very relevant application is given when estimating the credit risk of any company, as detailed in several research (Mongrut et al., 2011; Cerezo, 2012; Poveda, 2019; Virgilio et al., 2022). Pioneering authors who have used financial ratios to predict bankruptcy are Smith (1930), Fitzpatrick (1932), Smith & Winakor (1935), Merwin (1942), Chudson (1945), Beaver (1966) and Altman (1968). Several studies (Ahmeti & Zubanovic, 2020; Ariyo, 1986; Lundqvist & Strand, 2013) using financial indicators were aimed at improving the accuracy and applicability of models predicting bankruptcy (Tinoco & Wilson, 2013).

There are aspects that have not been thoroughly studied in bankruptcy estimation (Amendola et al., 2017) and this is why Ahmeti and Zubanovic (2020) and Wang et al. (2014) point out the importance of using statistical methods in the selection of appropriate financial ratios, to detect bankruptcy in different industries and time periods. Ariyo (1982) highlights the use of statistical techniques to select important financial variables in forecasting and Correa-Mejía and Lopera-Castaño (2020) state that the application of numerous financial variables can help to improve the accuracy of statistical models.

Different categories have been used to classify financial ratios (Correa et al., 2019). Beaver's (1966) financial ratios have been widely used in the financial management and profitability of companies (Charpentier et al, 2013) From an initial list of 22 variables, Beaver (1966) classified them into five standardized accounting categories (liquidity, profitability, leverage, solvency and activity) (Aquino, 2010; Valladares & Flores, 2005). Evaluating these aspects provides crucial information to investors and analysts, as well as to entrepreneurs to make a more effective management of their business (Haro et al., 2023).

Beaver (1966) examined some predictions regarding the mean values, for 79 healthy and 79 failed companies of six financial variables: cash flow over total liabilities, net income over total assets, total liabilities over total assets, working capital over total assets, current assets over current liabilities and non-credit interval (current assets minus inventory minus current liabilities) (Aquino, 2010; Camacho & Gómez, 2019).

Based on the calculation of the averages of the financial ratios in the five years prior to bankruptcy, Beaver (1966) concluded that the companies that had financial fragility with respect to healthy companies had lower cash flow, lower current asset reserves, greater dependence on contracting financial obligations and less willingness to meet the obligations already contracted

(García, 2019). The cash/total liabilities variable, called Beaver's ratio, is the indicator with the highest predictive capacity of business failure in Beaver's model (Somoza & Vallerdú, 2009). The firm showing an effective/liability ratio below 0.30 is determined as a failed firm (Camacho et al., 2019).

The cash/liability ratio had an error of 13% in the first year of bankruptcy anticipation, and 22% in the fifth year prior to bankruptcy (Aquino, 2010). The second best predictor was the ratio of total income to total assets, followed by the ratio of total assets to total liabilities, then the ratio of working capital to total assets, the current ratio and lastly the no credit interval (with an error of 23% for the first year of anticipation and 37% errors five years prior to failure) (Aquino, 2010).

Later, Liang et al. (2016), established nine categories for financial ratios, taking the categories of Beaver (1966) and following Fedorova et al. (2013), added three novel categories (growth structure, ownership structure and retention of key human capital) (Correa et al., 2020). Other studies classify financial ratios into three categories: liquidity, profitability and debt (Ochoa et al., 2009; Lopez et al., 2015).

According to Gupta et al. (2018) liquidity is a short-term variable that firms have to meet their immediate obligations. Liquidity has a fundamental role in the development of organizations, due to the fact that it involves substantial elements such as the fulfillment of obligations within the established deadlines, the acquisition of products and services with better prices and more efficient negotiations (Solórzano-Quito & Vásconez-Acuña, 2021).

Lack of liquidity can lead to a decrease in current assets, which could affect the company's ability to continue its operations and meet its obligations (Simbaña, 2020). In addition, the lack of liquidity may cause the establishment of survival strategies such as cost reduction, exploration of new sources of financing or negotiation with creditors, which may affect the solvency and profitability of the corporation (Terreno et al., 2020).

Studies have shown that liquidity management has an impact on company performance and companies with a higher level of liquid reserves have a lower risk of default, which reduces the probability of going bankrupt (Li & Xia, 2014). In addition, working capital management is to enable a company to maximize profits while preserving liquidity (Ramirez, 2014). However, increasing profits by sacrificing liquidity can lead to financial fragility, so it is important to have an optimal balance between profitability and liquidity (Panigrahi, 2014).

Shahdadi et al. (2020) states that intellectual capital can have a mitigating effect on the liquidity ratio and the probability of bankruptcy. In addition, it has been observed that long-term indebtedness has a negative relationship with liquidity, companies with long-term debts tend to have lower current capital and will have the need to reduce current assets to offset the cost of liabilities, which would increase the chances of going bankrupt (Thakor, 2018).

The current ratio can be used to measure the liquidity of a firm. Several studies (Nindita et al., 2014; Hidayat & Meiranto , 2014) show that the liquidity ratio has a significant negative impact on financial stress. On the contrary, Adiyanto (2021) and Septiani & Dana (2019) find a significant positive effect between the liquidity ratio and financial stress; on the other hand, the liquidity ratio has a significant positive effect on financial stress. in other research (Dirman (2020), Agustini & Wirawati (2019), Wulandari & Fitria (2019) and Sucipto & Muazaroh (2017)).

Jabeaur et al (2018) states that profitability is a business aspect that denotes the sustainability of the firm in the long term. Through the financial profitability ratios, comparisons are made between the amount invested by shareholders and the company, and the profits generated in the exploitation of the corporate purpose (Jayasekera, 2018). Bredart et al (2018) affirms that profitability allows knowing the creation of future cash flows. Correa et al (2020) points out that profitability and liquidity converge over time. Return on assets has been used in several research studies to measure profitability. The ROA indicator has a significant positive impact on business failure (Agustini & Wirawati (2019), Wulandari & Fitria (2019), and Sucipto & Muazaroh (2017)); but in other studies conducted by Giovanni et al, (2020) and Dirman (2020), there is evidence that ROA has a significant positive impact on failure; there are different results indicating that ROA has no effect on business failure (Nindita et al., (2014) and Hidayat & Meiranto (2014)).

Leverage measures the extent to which the firm's assets are financed with debt ((Fahmi, 2020) The liability to equity ratio is widely used to measure leverage. Giovanni et al., (2020), Septiani & Dana (2019), and Nindita et al., (2014) indicate that leverage has a negative effect on bankruptcy; other studies Agustini & Wirawati (2019) and Hidayat & Meiranto (2014) indicate that leverage has a positive impact on bankruptcy and other studies Dirman (2020), Wulandari & Fitria (2019), and Sucipto & Muazaroh (2017) state that leverage has no effect on bankruptcy.

Activity ratios are responsible for measuring the degree of effectiveness of the firm in employing its resources (Arifiana et al., 2022). Total active turnover is an example of this type of indicator. Studies by Agustini & Wirawati (2019) and Hidayat & Meiranto (2014) state that the activity ratio has a negative effect on fragility; other authors Sucipto & Muazaroh (2017) state that the activity ratio has no effect on fragility.

According to Ortiz (2018), the use of financial indicators is the most effective way to analyze the profitability of companies. These indicators allow mathematically comparing different financial accounts of the financial statements to evaluate the performance of the organization in different areas. Coral and Gudiño (2013) also agree that financial indicators are a fundamental tool for financial analysis, since they allow measuring and evaluating the behavior of financial accounts in different aspects of the company. Financial indicators are a key tool for financial analysts and managers, providing them with valuable information on the financial health of the company.

- Bankruptcy prediction models

For the estimation of insolvency states, the specialized literature focuses on three extensive developments: the application of multivariate statistics (Altman, 1968), logistic regressions (Ohlson, 1980) and data mining (Barboza et al., 2017). Each line of research has had a breakthrough and has been massively employed in many articles using different specifications, countries, and periods (Parra et al., 2022). The creation of reliable business failure forecasting models can be of great use to investors, creditors, managers, auditors, and regulatory agencies, as it allows them to make well-informed decisions and protect themselves from financial losses (Sunarjanto et al., (2016).

- Multivariate statistics technique.

Altman (1968) developed an indicator known as Z-score, which is a tool for predicting the probability of bankruptcy in companies. This model is based on financial and accounting information and considers the market value of companies as an important factor. In addition, the model includes two thresholds: one with a high probability of bankruptcy and another with a low probability of bankruptcy, with a margin of error that allows distinguishing between companies at risk and those that are not. It should be noted that this model has also been used for budgetary fines, such as the detection of tax evasion (Moreno and Bravo 2019). This model comprised five financial variables: working capital to total assets; retained earnings to total assets; earnings before interest and taxes to total assets; market value of equity to total liabilities and sales to total assets; these variables indicate the probability of bankruptcy if Z<1.81 indicates a high probability of bankruptcy, a Z between 1.81 and 2.99 is in the gray zone and a Z>2.99 indicates a low probability of bankruptcy (León and Vargas 2023).

- Logistic Regression.

According to Ruiz (2015), financial insolvency represents one of the biggest challenges for companies. In his research study, **the logistic regression** model was used to predict at least one year in advance the bankruptcy of small and medium-size companies in Colombia. The sample consisted of 838 active MSMEs and 57 in liquidation, and the indicators used were: working capital/total assets, sales/total assets, and gross margin. The model was based on Altman's Z1 model and the CA-Score, and the results indicated that the best model for predicting financial insolvency is Altman's Z1 (León and Vargas 2023).

The logistic regression model was first used by Ohlson (1980) to predict firm bankruptcy. In his study, Ohlson used an unbalanced data set and a large sample of bankrupt firms, collecting financial data over the period 1970 to 1976. He calculated nine different ratios and constructed a logistic regression model to predict the probability of corporate bankruptcy. Similarly, in his study, du Jardin (2015) examined the performance of a logistic regression model, as well as other models, over a period of up to three years prior to bankruptcy using data from French companies. To achieve this, the author used the underlying data and grouped the firms into different bankruptcy processes. The results obtained showed that the logistic regression model provided better long-term predictions compared to other commonly used tools. This research demonstrates the effectiveness of traditional bankruptcy prediction models over a three-year horizon.

In his study Olhson (1980) used the econometric methodology of **conditional logit analysis**, together with multivariate analysis or also known as discriminant analysis (MDA). The objective of this model was to obtain data from three years prior to the date of failure, analyzing the balance sheet, income statement, funds statement and accounting report, for a sample of 105 companies that participated in the stock market and failed between the years 1970 and 1976. In addition to this sample, information from 2058 solvent rated companies participating in the market was included. Ohlson selected nine ratios and pointed out that special attention should be paid to four elements that could lead to failure, among which company size, financial structure capacity, profitability and short-term liquidity stand out (Dussan and Briceida 2021).

In another study by Viana (2021), the logistic regression model was used to predict the risk of bankruptcy in Colombian hospitals. The financial indicator used as a dependent variable was ROE, since it shows the relationship between profit and equity. In addition, 14 financial ratios were used as independent variables, among which the Debt-Equity, return on assets and current asset turnover ratios stand out, which are explanatory of the dependent variable and are therefore definitive when anticipating the insolvency of a company. In summary, the logistic regression model used in this study allowed predicting the risk of bankruptcy in Colombian hospitals effectively, using key financial indicators as dependent and independent variables (León and Vargas 2023).

- Machine learning models.

In contrast to traditional models such as the Altman Z-score, **machine learning models** have proven to have a greater predictive capacity in the prediction of corporate insolvency. Vinet & Zhedanov (2011) defines machine learning model as "the study of algorithms that allow machines to improve their performance on a specific task as they are provided with new data" (p.217)

Dussan and Briceida 2021). Sandoval (2018) states that devices that are trained to learn and understand acquire skills comparable to those of humans, as they possess a greater capacity and speed of development than any person (Dussan and Briceida 2021).

Among machine learning models, the random forest stands out, which is a machine learning algorithm registered by Leo Breiman and Adele Cutler (2001), in which the output of several decision trees is combined to reach the expected result; this algorithm allows handling classification and regression problems (IBM s.f.). In addition, Gómez (2020) defines random forests as an analytical tool used to investigate the failure of companies. These random forests are considered as a form of supervised artificial intelligence, in which a sample is evaluated through attributes and the results of operations are displayed through branches.

To check the validity of the random forest, it is necessary to assess the quality of the tree structure through cross-validation. This technique involves the division of the sample into subsamples. Cross-validation produces a single, final tree model. The risk estimate by cross-validation for the final tree is calculated as the average of the risks of all trees (IBM s.f.). According to DataScientest (2022) the number of trees parameter is usually adjusted by cross-validation. This technique of evaluating Machine Learning algorithms involves training and testing the model on fragments of the original data set. In conclusion, cross-validation is a useful tool to ensure that the Machine Learning model is accurate and generalizable to new datasets by adjusting the number of trees through cross-validation, the accuracy of the model can be further improved, and overfitting can be avoided.

Correa et al., (2020) used financial indicators to forecast corporate insolvency one year in advance and concluded that financial indicators are useful for predicting corporate insolvency, but non-traditional methodologies such as the boosting algorithm that consider information asymmetry should be used. Franco (2022). found that XGBoost, SVM, Smote, RFY DT algorithms present a much higher predictive capacity than traditional methodologies, focused on a time horizon before the event given their higher accuracy. Overall, machine learning models appear to be an effective alternative to address the problem of corporate insolvency. This study makes a comparison between logistic and random forest algorithms.

3. Empirical methodology

The binary logistic regression model was used for data analysis, which is a useful tool for predicting the probability of a binary event, such as success or failure, as a function of the independent variables, and logistic regression coefficients can be used to estimate the likelihood ratio of each independent variable in the model (IBM 2021). In this article, I will address the issue of estimating the probability that a company belongs to a certain class, specifically, whether it is bankrupt or not. This classification is of vital importance to investors and financial analysts, as it allows them to assess the risk associated with a company and make informed decisions. To carry out this estimation, a model will be used that analyzes different variables and characteristics of the company, such as its financial history, liquidity indicators, profitability and indebtedness, among others. These variables are used as inputs to the model, which then calculates the probability that the company belongs to the class of bankrupt companies or to the class of non-bankrupt companies. Once the probability is obtained, a threshold of 50% is set. If the calculated probability is greater than 50%, the model predicts that the company belongs to the class of bankrupt companies. (Correa et al., (2020))

This means that there is a high probability that the company is facing financial difficulties and may declare bankruptcy in the near future. On the other hand, if the probability is less than 50%, the model predicts that the company does not belong to the class of bankrupt companies. In this case, the company is considered to be solvent and in a stable financial situation. This implies that the company has a low probability of facing serious financial problems in the short term. It is important to note that this probability estimate is not infallible and may be subject to errors. However, it is considered a useful tool for assessing a company's financial risk and making informed investment and financial analysis decisions. (IBM s.f.)

The logistic regression technique is a statistical method used to predict the probability of occurrence of a binary event, i.e. an event that can have only two possible outcomes, such as yes or no, success or failure, etc. The sigmoid function, also known as logistic function, is used to transform the linear regression model to fit the data. The sigmoid function is an S-shaped curve that has the property that the output values are constrained between 0 and 1. This is important in the context of logistic regression, since it allows us to interpret the results as probabilities. For example, if the function predicts a value of 0.8 for a particular event, it can be interpreted as an 80% probability that the event will occur. The logistic function is mathematically defined as: (DataScientest (2022))

$$F(z_i) = \frac{1}{1 + e^{-z_i}}$$

A random variable z is considered, and the exponential function is used to estimate the Logistic Regression model (Brooks , 2019). The estimated model is represented as:

$$P_i = \frac{1}{1 + e^{-(\beta_1 + \beta_2\, x_{2i} + \cdots + \beta_k\, x_{ki} + u_i)}}$$

where Pi denotes the probability that the variable will take the value of one (γi = 1). (DataScientest (2022).

On the other hand, the Random Forest is a set of decision trees that looks for the best feature to split the nodes, resulting in a greater diversity of trees and better model performance. It is used to predict whether a company is bankrupt or healthy and also calculates an importance score for each variable, indicating its contribution to the ranking. (Dussan and Briceida 2021)

Emphasizing that misclassifying an unhealthy company as healthy incurs greater costs than misclassifying a healthy company as unhealthy. The relative importance of one or the other error depends on the user's opinion, but for the purposes of this paper. The purpose is to reduce Type I error due to the higher costs involved. In general terms, the relevance of reducing classification errors and strengthening the consistency of the results in this context is highlighted. (IBM s.f.)

Following the same line, performance measures were implemented by using a confusion matrix as a model validation method. This matrix provides an overview of how many companies were incorrectly classified as solvent or bankrupt, as well as how many companies were misclassified as solvent or bankrupt. Considering that the objective is to reduce the Type 1 Error, it is desirable to maximize the sensitivity or true positive rate (TPR), which is calculated as follows (DataScientest (2022):

$$sensitivity = \frac{TN}{(TN + FP)}$$

Where TP is True Positive FN is False Negative. True Positives are those cases in which the company was correctly identified as bankrupt, which means that adequate measures were taken to minimize losses and protect investors' interests. On the other hand, False Negatives are those cases where the company was classified as healthy but was actually bankrupt. This can lead investors to make wrong decisions and suffer significant losses. (Dussan and Briceida 2021)

In addition, it is essential to consider False Positive Rates (FPR), which are determined as follows:

$$specificity = \frac{TN}{(TN + FP)}$$

In the context of True Negative (TN) and False Positive (FP), TN represents cases where healthy companies are correctly classified as such. This means that the model or classification system has correctly identified that a company is financially healthy and classified it as such. On the other hand, FP refers to cases where healthy companies are incorrectly classified as failing. This implies that the model or classification system has made a mistake in identifying a healthy company as being in financial trouble or bankrupt. These concepts are particularly relevant in the area of financial risk assessment and the classification of companies based on their financial health. The ability of a model or rating system to minimize FP and maximize TN is critical to ensure accurate and reliable assessment of companies. (IBM s.f.)

Finally, I proceed to calculate the ROC curve's area under the curve (AUC). In this case, the ROC curve serves to plot the ratio of true positives against false positives (Géron, 2022). A ROC AUC score greater than 0.5 is necessary to consider acceptance of the model, as 0.5 indicates a random guess. The predictive power and accuracy of the model are directly proportional to a higher ROC AUC score, with a score closer to 1 representing both stronger predictive capabilities and accuracy (Barboza et al., 2017).

4. Results

The first models that were estimated from the data collected correspond to logistic models. Although the resampling techniques (up, down, SMOTE, ROSE) were used for illustrative purposes, the estimation was also done using the training subsample without resampling. The results of the latter estimation are shown in Table ## and reflect that indeed, since there is no balance in the sample, the model loses all capacity to predict business failure (see Table 1).

Table 1. Logistic model prediction performance - Data without resampling

		Actual Values		Metric	
		0	1	Accuracy	0.9832
Predicted Values	0	3212	52	Sensitivity	0.00
	1	3	0	Specificity	0.999

Because of this, it is important to address the problem of imbalance in the samples to run forecasting models, as well as it is essential to analyze different prediction metrics as a whole, because as observed in this particular case, a specification of 100% and an accuracy rate of 98.20% were obtained, without this implying that the model has a good performance; in fact, the model obtains a sensitivity of 0.00%, reflecting that it is not able to identify any insolvent company. It is precisely this problem that is controlled by using resampling methods (upsampling, downsampling, random oversampling examples (ROSE) and synthetic minority oversampling technique (SMOTE)).

Continuing with the estimations made from the different resampling techniques, Table 3, which summarizes the prediction techniques for each of the estimated models, is presented. In the first instance, it is observed that with all models there is a substantial improvement in the metrics in relation to the models without balancing, with accuracy rates exceeding 60% (except in one case), which when reviewed together with other metrics, give us indications of improvements in the predictions compared to the models previously described.

Table 2. Random forest model, prediction performance - Data without resampling

		Actual Values		Metric	
		0	1	Accuracy	0.9841
Predicted Values	0	3220	52	Sensitivity	0.00
	1	0	0	Specificity	1.00

Table 3. Summary of prediction results for each type of resampling.

Metric	Up	Down	ROSE	SMOTE
		Logit model		
Accuracy	0.7995	0.7995	0.802	0.8331
Sensitivity	0.73077	0.78846	0.76923	0.61538
Specificity	0.80062	0.79969	0.80248	0.83665
		Random forest		
Accuracy	0.9838	0.761	0.0538	0.8285
Sensitivity	0	0.80769	1	0.61538
Specificity	0.9996	0.76025	0.03851	0.83199

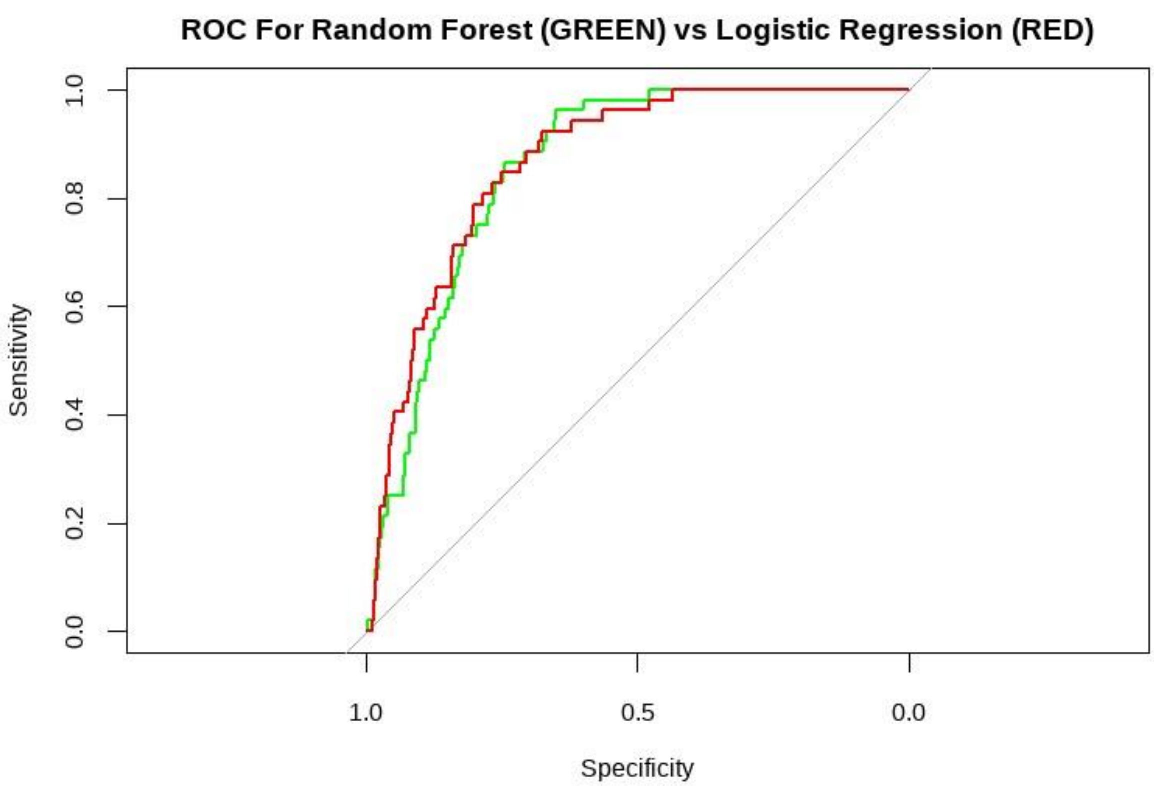
Source: own source.

As for the logistic models, whose coefficients are recorded in Table A7, 3, it is observed that although they achieve a sensitivity rate above 70% in three of the four resampling classes, the correct identification of non-insolvent companies is high in relation to other models (not in upsampling). In particular, the estimation that achieves the best balance between sensitivity and specificity is achieved with the undersampling technique.

In this case, the model classifies 41 companies as insolvent and 2575 as non-insolvent. Of the 3267 companies in the base, the model had an accuracy rate of 79.95% (2616), surpassing the rate achieved by the random forest model. The improvement is mainly due to a better detection of non-insolvent companies (reflected in the specificity rate), which in turn translates into a reduction in the false positive rate (20.03%). Of the 14 variables that estimate the model, those with the highest statistical significance are cash flow to liabilities, current assets to assets, liabilities to equity, equity to liabilities, gross margin, net margin, return on equity (roe).

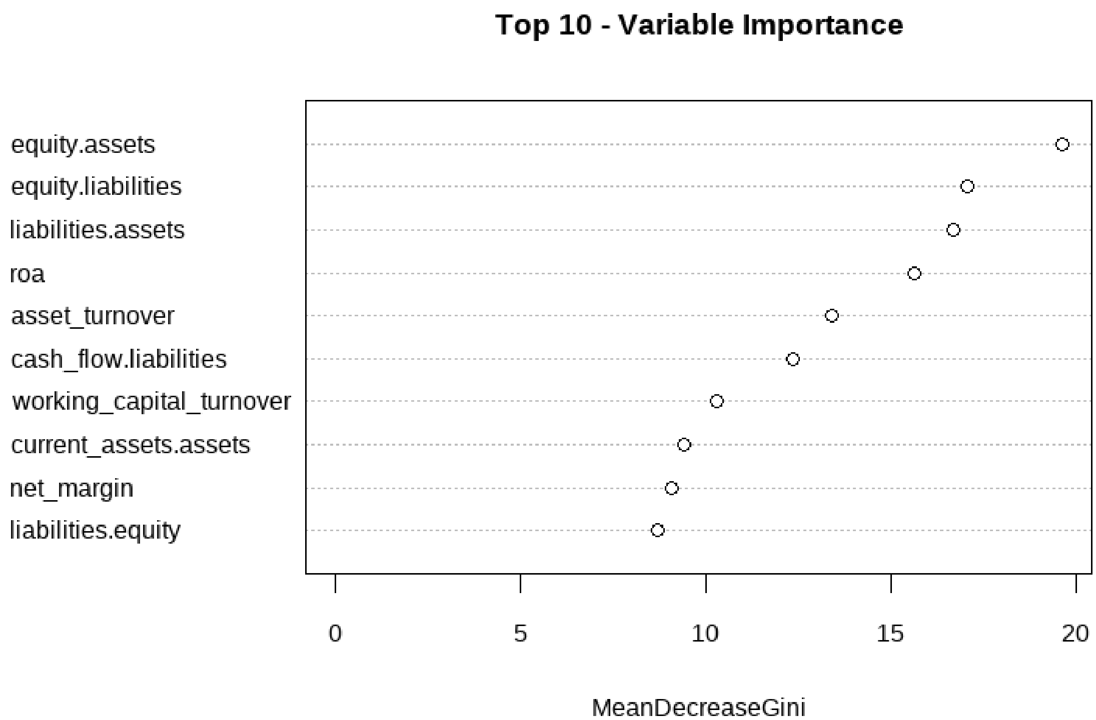
The best performance of the logit model can be seen graphically in Figure 1, where the ROC curve of this model is always above the estimates made with the other models. This shows that the relationships between the variables captured by the logit model generate a better prediction of business failure. However, the random forest model should not be completely discarded, since this model captures non-linear interactions between predictors and establishes complex relationships between the independent variables, this type of non-linear relationships is not evident in conventional models.

Figure 1 ROC Curve.



Source: Own source. [1] "Accuracy % of random forest: 0.763447432762836" # [1] "Accuracy % of logistic regression: 0.799511002444988". Area under curve of random forest: 0.857946129956999". Area under curve of logistic regression: 0.866435738174869

Figuer 2 : Random Foreat Variable Importance.



Now, although under this method there is no specification of a global model that can be applied to , it is possible to determine which are the most important predictors. These are shown in Figure 2, which highlights the variables corresponding to the equity to assets, equity to liabilities, liabilities to asses, roa, asset turnover, cash flow to liabilities, working capital turnover, current assets to assets, net margin and liabilities to equity.

5. Conclusion

This paper carries out a comparison of the quality of business failure prediction, derived from the estimates of a logistic model and random forests. Based on the estimates it can be inferred, first, that although the logistic model applied to Colombian firms in the industrial manufacturing sector generates a high specificity rate, the model is not able to identify any insolvent firm.

Because the classification and prediction models are very sensitive to the presence of unbalanced data that have biases in their predictions, it was necessary to resort to the use of resampling techniques to generate balanced samples. Specifically, the undersampling resampling technique generates the best performance metrics in business failure forecasting by using estimation techniques and making the respective forecasts on the test subsample.

Among the estimated models, the random forest method slightly improves the detection of insolvent firms compared to the logit model, but the accuracy rate is lower in most of the resampling cases. Overall, the logit model slightly outperforms the random forest model in its prediction performance. The importance of the random forest lies in the fact that it is possible to find complex non-linear relationships in the determinants of corporate insolvency, which conventional models cannot detect. This is a a very important advantage for machine learning models.

Finally, it is worth mentioning that, according to the random forest model, the variables that have a greater incidence in the explanation of business failure are: equity to assets, equity to liabilities, liabilities to asses, roa, asset turnover, cash flow to liabilities, working capital turnover, current assets to assets, net margin and liabilities to equity. The logit mpdel variables with the highest statistical significance are cash flow to liabilities, current assets to assets, liabilities to equity, equity to liabilities, gross margin, net margin, return on equity (roe).

Appendix

Table A1. Variables and definition

	Categoty	Variable	Calculation
0	class	class	categorical variable that establishes the class of the firm between solvent (class=0) and insolvent (class=1).
1	Liquidity	current_ratio	Total current assets/Total current liabilities

2	Liquidity	acid_test_ratio	(Total current assets-Current inventories-Non-current inventories)/Total current liabilities
3	Liquidity	working_capital/assets	Working capital/Total assets
4	Liquidity	cash/current_liabilities	Cash and cash equivalents/Total current liabilities
5	Liquidity	current_assets/assets	Total current assets/Total assets
6	Leverage	cash_flow/liabilities	Cash and cash equivalents at the end of the period/Total liabilities
7	Leverage	liabilities/equity	Total liabilities/Total equity
8	Leverage	equity/assets	Total equity/Total assets
9	Leverage	equity/liabilities	Total equity/Total liabilities
10	Leverage	operating_profit/financial_costs	Profit from operating activities/financial costs
11	Leverage	liabilities/assets	Total liabilities/Total assets
12	Leverage	sales/current_assets	Income from ordinary activities/Total current assets
13	Profitability	EBIT/sales	EBIT/Income from ordinary activities
14	Profitability	EBITDA/sales	EBITDA/Income from ordinary activities
15	Profitability	gross_margin	Gross profit/Income from ordinary activities
16	Profitability	operating_margin	Profit from operating activities/Income from ordinary activities
17	Profitability	net_margin	Profit/Income from ordinary activities
18	Profitability	roe	Profit/Total equity
19	Profitability	roa	Profit/Total assets
20	Activity/Operating	inventory turnover	Income from ordinary activities/average inventory
21	Activity/Operating	assets turnover	Income from ordinary activities/average total assets
22	Activity/Operating	working capital turnover	Income from ordinary activities/average working capital

*Notes: working_capital= Total current assets-Total current liabilities.
Cash and cash equivalents: Proxy of cash.
Cash and cash equivalents at the end of the period: Proxy of cash flow.
Profit from operating activities: Proxy of operating
profit.
Income from ordinary activities: Proxy of sales.
EBIT: Profit from operating activities.
EBITDA: Profit from operating activities+Adjustments for depreciation and amortization expense+Adjustments for provisions.
Profit: Proxy of net profit.
Source: own elaboration.

Table A2. Descriptive statistics -imputed data- NON-INSOLVENT COMPANIES

described_variables	na	mean	sd	skewness	minimum	p25	median	p75	maximum
acid_test_ratio	0	84.12	3625.1	71.45	-231.29	0.68	1.07	1.78	361869.25
asset_turnover	0	1.21	1.01	3.81	0	0.63	1.01	1.52	21.62
cash.current_liabilities	0	2.56	62.67	40.17	0	0.04	0.12	0.37	4490.43
cash_flow.liabilities	0	0.55	11.36	58.88	0	0.02	0.07	0.23	1012.47
current_assets.assets	0	0.59	0.25	-0.26	0	0.41	0.6	0.79	1

current_ratio	0	80.79	3431.6	79.44	0	1.19	1.7	2.64	361869.25
ebit.ssales	0	971.68	7859.9	8.81	-60404	0.03	0.07	0.12	235744
ebitda.sales	0	960.99	7891.2	9.99	-60404	0.05	0.09	0.15	290324
equity.assets	0	0.35	6.46	-114.3	-788.73	0.3	0.47	0.64	1
equity.liabilities	0	2.9	31.45	30.01	-1	0.43	0.88	1.77	1479.04
gross_margin	0	-29.98	267.42	-14.04	-12869	0.18	0.26	0.37	1
inventory._turnover	0	19.42	327.29	113.69	0	3.19	5.77	12.13	39980.1
liabilities.assets	0	0.65	6.46	114.3	0	0.36	0.53	0.7	789.73
liabilities.equity	0	2.21	41.21	-23.36	-3471.22	0.48	1.04	2.02	2280.32
net_margin	0	141.95	1865.6	4.75	-108252	0.01	0.03	0.07	107564
operating_margin	0	492.67	4359.1	15.72	-60404	0.03	0.07	0.12	235744
roa	0	-0.02	5.86	-126.45	-742.83	0	0.03	0.08	2.03
roe	0	0.05	7.15	10.18	-409.36	0.02	0.08	0.18	574.43
sales.current_assets	0	3.82	98.21	104.41	0	1.25	1.87	2.73	11552.68
working_capital.assets	0	0.17	6.23	-125.92	-788.73	0.07	0.23	0.4	1
working_capital_turnover	0	17.67	1143	114.98	-7233.24	1.37	3.29	6.7	140045

*Notes: na: missing values, sd: standard deviation, p25: 25th percentile, p75: 75th percentile.
Source:own elaboration.

Table A3. Descriptive statistics -imputed data- INSOLVENT COMPANIES

described_variables	na	mean	sd	skewness	minimum	p25	median	p75	maximum
acid_test_ratio	0	1.99	4.98	7.04	-0.9	0.41	0.8	1.69	53.03
asset_turnover	0	0.76	0.69	2.09	0	0.26	0.62	0.96	4.6
cash.current_liabilities	0	0.28	0.98	8.75	0	0.01	0.05	0.16	12.45
cash_flow.liabilities	0	0.04	0.06	3.14	0	0	0.02	0.05	0.41
current_assets.assets	0	0.5	0.25	0.1	0.01	0.31	0.48	0.69	0.99
current_ratio	0	3.53	8.61	5.87	0.02	0.76	1.39	2.68	70.81
ebit.ssales	0	887.98	7222.39	7.99	-1373.68	-0.14	0.01	0.06	61608.93
ebitda.sales	0	853.86	6928.5	8.02	-638.47	-0.08	0.03	0.09	61769.42
equity.assets	0	0	0.65	-3.31	-4.39	-0.02	0.19	0.34	0.81
equity.liabilities	0	0.27	0.58	1.89	-0.81	-0.02	0.23	0.51	4.21
gross_margin	0	-35.14	291.09	-8.46	-2745.11	0.11	0.24	0.36	1
inventory._turnover	0	9.33	18.94	6.2	0	1.63	3.37	9.15	209.67
liabilities.assets	0	1	0.65	3.31	0.19	0.66	0.81	1.02	5.39
liabilities.equity	0	2.43	13.49	1.05	-66.93	-1.58	1.89	4.01	85.99
net_margin	0	26.85	578.1	10.13	-2186.36	-0.18	-0.01	0.03	7870.09
operating_margin	0	389.36	3226.63	8.26	-1373.68	-0.14	0.01	0.06	32547.51
roa	0	-0.1	0.51	-8.86	-5.92	-0.08	-0.01	0.02	0.27
described_variables	na	mean	sd	skewness	minimum	p25	median	p75	maximum
roe	0	-0.03	1.27	-3.4	-12.11	-0.11	0.02	0.19	6.93
sales.current_assets	0	1.86	1.71	2.13	0	0.71	1.42	2.37	11.89
working_capital.assets	0	0.04	0.51	-3.54	-4.53	-0.12	0.11	0.34	0.84
working_capital_turnover	0	0.22	24.97	-1.15	-214.02	-1.12	1.12	3.34	211.44

*Notes: na: missing values, sd: standard deviation, p25: 25th percentile, p75: 75th percentile.
Source: own elaboration.

Table A4. Descriptive statistics -standarized data- NON-INSOLVENT COMPANIES

described_variables	na	mean	sd	skewness	minimum	p25	median	p75	maximum
acid_test_ratio	0	0	1	-50.15	-99.65	-0	-0.03	-0	41.29
asset_turnover	0	0.01	1	-0.02	-2.28	-0.6	0.03	0.63	4.19
cash.current_liabilities	0	0.01	1	0.63	-1.19	-0.9	-0.26	0.76	2.14
cash_flow.liabilities	0	0	1	58.88	-0.05	-0.1	-0.04	-0	89.77
current_assets.assets	0	0.01	1	-0.09	-2.17	-0.8	0.01	0.81	1.79
current_ratio	0	0	1	-0.01	-3.59	-0.6	-0.05	0.58	3
ebit.ssales	0	0	1	-40.17	-89.03	-0.1	-0.08	-0.1	18.02
ebitda.sales	0	0	1	-39.84	-88.98	-0.1	-0.08	-0.1	21.64
equity.assets	0	0.02	1	0.13	-3.96	-0.7	-0.08	0.66	2.74
equity.liabilities	0	0.02	1	-0.19	-5.59	-0.6	-0.06	0.58	3.86
gross_margin	0	0	1	0.76	-3.7	-0.5	-0.15	0.34	4.81
inventory._turnover	0	0.01	1	-0.02	-2.8	-0.6	-0.03	0.64	3.76
liabilities.assets	0	-0.02	1	-0.13	-3.02	-0.6	0.1	0.63	4.54
liabilities.equity	0	0	1	-3.29	-73.69	-0.1	-0.03	-0	63.77
net_margin	0	0	1	-29.83	-86.14	-0.1	-0.06	-0.1	46.19
operating_margin	0	0	1	-37.15	-88.17	-0.1	-0.08	-0.1	32.87
roa	0	0.01	1	5.85	-15.75	-0.3	-0.04	0.29	39.6
roe	0	0	1	11.91	-57.13	-0	0	0.02	81.84
sales.current_assets	0	0.01	1	-0.09	-2.72	-0.5	0.03	0.56	5.81
working_capital.assets	0	0.01	1	0.32	-10.32	-0.7	-0.13	0.54	3.7
working_capital_turnover	0	0	1	13.11	-64.32	0.01	0.02	0.03	83.38

*Notes: na: missing values, sd: standard deviation, p25: 25th percentile, p75: 75th percentile.
Source:own elaboration.

Table A5. Descriptive statistics -standarized data- INSOLVENT COMPANIES

described_variables	na	mean	sd	skewness	minimum	p25	median	p75	maximum
acid_test_ratio	0	-0.02	0.1	4.13	-0.1	-0.1	-0.03	-0	0.43
asset_turnover	0	-0.62	1	0.38	-2.28	-1.4	-0.61	-0	2.34
cash.current_liabilities	0	-0.42	0.9	1.43	-1.19	-1.1	-0.77	-0	2.14
cash_flow.liabilities	0	-0.04	0	3.14	-0.05	-0.1	-0.05	-0	-0.01
current_assets.assets	0	-0.36	1	0.28	-2.13	-1.1	-0.48	0.37	1.75

current_ratio	0	-0.28	1.3	0.13	-3.49	-1.2	-0.35	0.6	2.76
ebit.ssales	0	-0.01	0.7	7.82	-1.28	-0.1	-0.08	-0.1	5.61
ebitda.sales	0	-0.01	0.7	7.98	-0.59	-0.1	-0.08	-0.1	5.64
equity.assets	0	-1.07	0.8	-0.19	-3.29	-1.5	-0.99	-0.6	1.54
equity.liabilities	0	-1.09	1	-0.97	-4.5	-1.4	-0.88	-0.5	1.42
gross_margin	0	-0.17	1.1	0.35	-3.7	-0.7	-0.25	0.28	4.81
inventory._turnover	0	-0.46	1.2	0.08	-2.8	-1.2	-0.55	0.39	2.48
liabilities.assets	0	1.06	0.9	0.58	-1.54	0.53	0.93	1.4	3.79
liabilities.equity	0	0.01	0.3	1.69	-1.59	-0.1	-0.01	0.04	2.2
net_margin	0	-0.05	0.3	7.25	-1.61	-0.1	-0.06	-0.1	3.59
operating_margin	0	-0.02	0.6	7.91	-1.4	-0.1	-0.08	-0.1	5.57
roa	0	-0.66	1.2	-4.11	-9.53	-0.8	-0.33	-0.2	2.04
roe	0	-0.01	0.2	-3.36	-1.71	-0	-0.01	0.02	0.97
sales.current_assets	0	-0.4	1.2	-0.12	-2.72	-1.2	-0.35	0.36	2.54
working_capital.assets	0	-0.47	1.1	-0.07	-4.74	-1.2	-0.51	0.28	2.73
working_capital_turnover	0	0	0.1	-7.85	-1.14	0.01	0.01	0.02	0.33

*Notes: na: missing values, sd: standard deviation, p25: 25th percentile, p75: 75th percentile.
Source:own elaboration.

Table A6. Correlation matrix.

Means, standard deviations, and correlations with confidence intervals

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. current_ratio	79.55	3403.84																
2. acid_test_ratio	82.80	3595.73	.99**															
			[.99, .99]															
3. working_capital.assets	0.17	6.18	.00	.00														
			[-.01, .02]	[-.01, .02]														
4. cash.current_liabilities	2.52	62.16	.39**	.40**	.00													
			[.38, .41]	[.39, .41]	[-.01, .02]													
5. cash_flow.liabilities	0.54	11.27	.36**	.41**	.00	.48**												
			[.34, .37]	[.40, .42]	[-.01, .02]	[.47, .49]												
6. current_assets.assets	0.59	0.25	.03**	.03**	.01	.03**	.04**											

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
			[.01, .04]	[.01, .04]	[-.00, .03]	[.01, .04]	[.02, .05]											
7. liabilities.equity	2.21	40.91	-.00	-.00	.00	-.00	-.00	.02**										
			[-.02, .01]	[-.02, .01]	[-.01, .02]	[-.02, .01]	[-.02, .01]	[.01, .04]										
8. equity.assets	0.35	6.41	-.01	-.00	.97**	-.03**	.00	-.02**	-.00									
			[-.03, .00]	[-.02, .01]	[.96, .97]	[-.05, -.02]	[-.01, .02]	[-.04, -.01]	[-.02, .01]									
9. equity.liabilities	2.86	31.20	.63**	.67**	.00	.39**	.63**	-.01	-.00	.01								
			[.62, .64]	[.66, .68]	[-.01, .02]	[.38, .41]	[.62, .64]	[-.02, .01]	[-.02, .01]	[-.01, .02]								
10. liabilities.assets	0.65	6.41	.01	.00	-.97**	.03**	-.00	.02**	.00	-1.00**	-.01							
			[-.00, .03]	[-.01, .02]	[-.97, -.96]	[.02, .05]	[-.02, .01]	[.01, .04]	[-.01, .02]	[-1.00, -1.00]	[-.02, .01]							
11. sales.current_assets	3.78	97.41	.01	.01	-.00	.01	.01	-.04**	-.00	-.01	.03**	.01						
			[-.01, .02]	[-.00, .03]	[-.02, .01]	[-.00, .03]	[-.00, .03]	[-.06, -.03]	[-.02, .01]	[-.02, .01]	[.01, .04]	[-.01, .02]						
12. ebit.ssales	970.33	7849.87	.09**	.09**	-.07**	.16**	.15**	-.05**	.01	-.08**	.23**	.08**	.01					
			[.08, .11]	[.08, .11]	[-.08, -.05]	[.14, .17]	[.14, .17]	[-.07, -.04]	[-.01, .02]	[-.10, -.07]	[.22, .25]	[.07, .10]	[-.01, .02]					

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
13. ebitda.sales	959.26	7876.41	.10**	.10**	-.06**	.17**	.16**	-.05**	.01	-.08**	.25**	.08**	.01	1.00**				
			[.08, .11]	[.08, .11]	[-.08, -.05]	[.15, .18]	[.15, .18]	[-.06, -.03]	[-.01, .02]	[-.10, -.06]	[.23, .26]	[.06, .10]	[-.01, .02]	[1.00, 1.00]				
14. gross_margin	-30.07	267.81	-.06**	-.06**	.04**	-.10**	-.12**	.05**	-.01	.05**	-.16**	-.05**	-.01	-.81**	-.78**			
			[-.07, -.04]	[-.07, -.04]	[.03, .06]	[-.11, -.08]	[-.14, -.11]	[.04, .07]	[-.02, .01]	[.03, .07]	[-.17, -.14]	[-.07, -.03]	[-.02, .01]	[-.81, -.80]	[-.78, -.77]			
15. operating_margin	491.00	4343.16	.10**	.10**	-.07**	.18**	.16**	-.04**	.01	-.09**	.25**	.09**	.01	.97**	.98**	-.66**		
			[.09, .12]	[.09, .12]	[-.08, -.05]	[.16, .19]	[.15, .18]	[-.05, -.02]	[-.01, .02]	[-.10, -.07]	[.24, .26]	[.07, .10]	[-.01, .02]	[.96, .97]	[.98, .98]	[-.67, -.65]		
16. net_margin	140.10	1852.02	.11**	.11**	-.06**	.18**	.15**	-.02**	.01	-.08**	.24**	.08**	.01	.74**	.77**	-.21**	.86**	
			[.09, .12]	[.09, .12]	[-.07, -.04]	[.17, .20]	[.14, .17]	[-.04, -.01]	[-.01, .02]	[-.09, -.06]	[.22, .25]	[.06, .09]	[-.00, .03]	[.73, .75]	[.77, .78]	[-.23, -.20]	[.86, .87]	
17. roe	0.05	7.09	-.00	-.00	-.00	-.00	-.00	-.02*	-.61**	-.00	-.00	.00	-.00	-.02**	-.02**	.02*	-.02**	-.01
			[-.02, .01]	[-.02, .01]	[-.02, .01]	[-.02, .01]	[-.02, .01]	[-.03, -.00]	[-.62, -.60]	[-.02, .01]	[-.02, .01]	[-.01, .02]	[-.02, .01]	[-.04, -.01]	[-.04, -.01]	[.00, .03]	[-.04, -.01]	[-.03, .00]
18. roa	-0.02	5.81	-.00	-.00	1.00**	-.00	-.00	-.01	.00	.97**	-.00	-.97**	-.00	-.07**	-.06**	.04**	-.07**	-.06**
			[-.02, .01]	[-.02, .01]	[1.00, 1.00]	[-.02, .02]	[-.02, .01]	[-.03, .00]	[-.01, .02]	[.96, .97]	[-.02, .02]	[-.97, -.96]	[-.02, .01]	[-.08, -.05]	[-.08, -.05]	[.03, .06]	[-.08, -.05]	[-.07, -.04]
19. inventory._turnover	19.26	324.65	.00	.00	.00	.00	.00	.01	.00	-.00	.01	.00	.00	.00	.00	.00	.00	.01
			[-.01, .02]	[-.01, .02]	[-.02, .02]	[-.01, .02]	[-.01, .02]	[-.01, .02]	[-.02, .02]	[-.02, .02]	[-.01, .02]	[-.02, .02]	[-.01, .02]	[-.01, .02]	[-.01, .02]	[-.01, .02]	[-.01, .02]	[-.01, .02]

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
20. asset_turnover	1.20	1.00	-.02**	-.02**	.01	-.03**	-.03**	.28**	.01	.01	-.06**	-.01	.02*	-.15**	-.15**	.14**	-.14**	-.09**
			[-.04, -.01]	[-.04, -.01]	[-.00, .03]	[-.05, -.02]	[-.04, -.01]	[.27, .29]	[-.01, .02]	[-.01, .02]	[-.07, -.04]	[-.02, .01]	[.00, .03]	[-.16, -.13]	[-.16, -.13]	[.12, .15]	[-.15, -.12]	[-.11, -.08]
21. working_capital_turnover	17.39	1133.78	-.00	-.00	-.00	-.00	-.00	-.00	.00	-.00	-.00	.00	.00	-.00	-.00	.00	-.00	-.00
			[-.02, .02]	[-.02, .02]	[-.02, .02]	[-.02, .01]	[-.02, .01]	[-.02, .01]	[-.01, .02]	[-.02, .02]	[-.02, .01]	[-.02, .02]	[-.01, .02]	[-.02, .01]	[-.02, .01]	[-.01, .02]	[-.02, .01]	[-.02, .02]

Note. *M* and *SD* are used to represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. The confidence interval is a plausible range of population correlations that could have caused the sample correlation (Cumming, 2014). * indicates $p < .05$. ** indicates $p < .01$.

Table A7. Logit model estimations for each type of resampling

	Dependent variable:				
	Bankruptcy (=1)	Bankruptcy (=1)		Bankruptcy (=1)	
	Without sampling	Up sampling	Down sampling	ROSE	SMOTE
current_ratio	-0.010	-0.061	0.452	0.162***	-0.168
	(0.127)	(0.040)	(0.369)	(0.021)	(0.185)
working_capital.assets	0.181	0.351***	0.205	-0.047**	0.409*
	(0.122)	(0.046)	(0.421)	(0.024)	(0.209)
cash.current_liabilities	0.459***	0.653***	0.313	-0.075***	0.965***
	(0.155)	(0.047)	(0.424)	(0.022)	(0.226)
cash_flow.liabilities	-89.260***	-85.799***	-68.969*	0.035	-120.546***
	(20.675)	(4.708)	(41.698)	(0.049)	(24.586)
current_assets.assets	-0.473***	-0.867***	-0.777**	-0.408***	-0.723***
	(0.118)	(0.041)	(0.327)	(0.022)	(0.179)
liabilities.equity	-0.065	-0.085***	-1.356***	-0.054*	-0.197
	(0.065)	(0.030)	(0.499)	(0.030)	(0.199)
equity.liabilities	-0.766***	-1.456***	-1.741***	-0.992***	-1.436***
	(0.073)	(0.027)	(0.245)	(0.022)	(0.123)
sales.current_assets	-0.218*	-0.655***	-0.460	-0.186***	-0.737***
	(0.113)	(0.044)	(0.363)	(0.022)	(0.186)
gross_margin	0.016	-0.198***	-0.336**	-0.103***	-0.480***
	(0.052)	(0.020)	(0.164)	(0.018)	(0.092)
net_margin	-0.060***	-1.172***	-1.977*	-0.306***	-2.440***
	(0.023)	(0.069)	(1.188)	(0.038)	(0.499)
roe	-0.061	-0.106***	-2.820**	0.018	0.009
	(0.049)	(0.028)	(1.322)	(0.023)	(0.586)
inventory._turnover	-0.108	0.114***	-0.063	-0.093***	0.190*
	(0.090)	(0.024)	(0.208)	(0.020)	(0.107)
asset_turnover	-0.035	-0.199***	-0.366	-0.344***	-0.248
	(0.138)	(0.044)	(0.366)	(0.024)	(0.194)
working_capital_turnover	0.005	0.019	-0.423	-0.018	-0.141
	(0.073)	(0.019)	(1.263)	(0.026)	(0.336)
Constant	-8.106***	-4.417***	-3.914**	-0.746***	-6.035***
	(0.840)	(0.182)	(1.605)	(0.025)	(0.958)

Observations	13,096	25,768	424	13,096	1,484
Log Likelihood	-907.233	-11,315.630	-174.955	-6,820.028	-620.636
Akaike Inf. Crit.	1,844.466	22,661.260	379.911	13,670.060	1,271.271

Note: Robust standard errors in parentheses. Significance is represented by *p<0.1; **p<0.05; ***p<0.01.

References

Adiyanto, Y. (2021). The Influence Of Institutional Ownership, Liquidity, And Company Size On Financial Distress: Empirical Study On Property & Real Estate Sub Sector Companies Listed On The Indonesia Stock Exchange 2015-2018. *International Journal Of Economics, Management, Business, And Social Science (Ijembis),* 1(1), 111-120. https://doi.org/10.59889/ijembis.v1i1.9

Agustini, N. W., & Wirawati, N. G. P. (2019). Pengaruh rasio keuangan pada financial distress perusahaan ritel yang terdaftar di bursa efek indonesia (BEI). *E-Jurnal Akuntansi*, *26*(1), 251-280. https://doi.org/10.24843/EJA.2019.v26.i01.p10

Ahmeti, L., & Zubanovic, A. (2020). The predictive power of financial ratios on bankruptcy: A quantitative study of non-listed limited liability SMEs companies in Sweden (Dissertation). Retrieved from https://urn.kb.se/resolve?urn=urn:nbn:se:hj:diva-48628

Alam, T. M., Shaukat, K., Mushtaq, M., Ali, Y., Khushi, M., Luo, S., & Wahab, A. (2021). Corporate bankruptcy prediction: An approach towards better corporate world. *The Computer Journal*, *64*(11), 1731-1746. https://doi.org/10.1093/comjnl/bxaa056

Alarcón, B. & Mejía, F. (2017). *Transparencia en los estados financieros y valor de la empresa: análisis de caso para microempresas colombianas.* Recuperado de: http://hdl.handle.net/10726/1650.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The journal of finance*, *23*(4), 589-609. https://doi.org/10.2307/2978933

Amendola, A., Giordano, F., Parrella, M. L., & Restaino, M. (2017). Variable selection in high-dimensional regression: a nonparametric procedure for business failure prediction. *Applied Stochastic Models in Business and Industry*, *33*(4), 355-368. https://doi.org/10.1002/asmb.2240

Andrade Pinelo, A. M. (2017). Ratios o razones financieras. http://hdl.handle.net/10757/622323

Aquino, S. (2010). Accounting indicators for credit risk analysis of firms: a historical perspective. *Economia Aziendale Online-*, *1*(2), 145-154. http://dx.doi.org/10.4485/ea2038-5498.145-154

Araica, Y.V., & Meda, C.E. (2017). Valoración Financiera de Empresas: Análisis de los estados financieros de Finca Nicaragua,S.A. para el periodo de diciembre 2014 y 2015. https://api.core.ac.uk/oai/oai:repositorio.unan.edu.ni:4859

Arifiana, R., & Khalifaturofi'ah, S. O. (2022). The effect of financial ratios in predicting financial distress in manufacturing companies. *Jurnal Riset Bisnis dan Manajemen*, *15*(2), 103-108. https://journal.unpas.ac.id/index.php/jrbm/article/view/5838/2445

Ariyo, A. (1982). *A SIGNAL DETECTION ANALYSIS OF AUDITORS'ANALYTICAL REVIEW JUDGMENTS*. The University of Arizona. https://www.proquest.com/openview/82dbcaedf64f7bce07202fc408ffde91/1?pq-origsite=gscholar&cbl=18750&diss=y

Ariyo, A. (1986). Financial Ratios for Bankruptcy Prediction A Consensus Approach. *Vikalpa*, 11(1), 47-54. https://doi.org/10.1177/0256090919860107

Atiya, A. F. (2001). Bankruptcy prediction for credit risk using neural networks: A survey and new results. *IEEE Transactions on neural networks*, *12*(4), 929-935. https://doi.org/10.1109/72.935101

Barboza, F., Kimura, H., & Altman, E. (2017). Machine learning models and bankruptcy prediction. Expert Systems with Applications, 83, 405-417. https://doi.org/10.1016/j.eswa.2017.04.006

Barboza, F., Basso, L. F. C., & Kimura, H. (2023). New metrics and approaches for predicting bankruptcy. *Communications in Statistics-Simulation and Computation*, *52*(6), 2615-2632.

Beaver, W. H. (1966). Financial Ratios As Predictors of Failure. *Journal of Accounting Research*, *4*, 71–111. https://doi.org/10.2307/2490171

Bernate, M. y Gómez, F. (2021). Predicción de la quiebra en las empresas. Una revisión de literatura. *Revista Activos*, 19(1). https://doi.org/10.15332/25005278.6684

BoneLLo, J., BrÉdart, X., & VeLLa, V. (2018). Machine learning models for predicting financial distress. *Journal of Research in Economics*, *2*(2), 174-185. https://doi.org/10.24954/JORE.2018.22

Breiman, L. (2001). Random forests. *Machine learning*, *45*, 5-32.
https://doi.org/10.1023/A:1010933404324

Brooks, C. (2019). Introductory econometrics for finance. Cambridge university press. https://www.cambridge.org/9781107661455

Camacho Laverde, N. J., & Gómez Silva, L. C. (2019). *Aplicación del Modelo Beaver en las pymes del sector comercio en Bogotá* (Doctoral dissertation, Corporación Universitaria Minuto de Dios). https://hdl.handle.net/10656/9876

Casanova Rubio, V. J. (2011). Valoración de los actuales métodos de previsión de insolvencia empresarial, así como su utilidad y precisión al introducir en ellos variables cualitativas que explican mejor la naturaleza de la empresa y su situación. http://hdl.handle.net/10251/14578

Cerezo, É. C. (2012). Metodología para la medición del riesgo de crédito de empresas no financieras (Doctoral dissertation, Universitat de Barcelona). https://dialnet.unirioja.es/servlet/tesis?codigo=178051

Charpentier, J. A. V., Gómez, M. B., & Rojas, J. M. C. (2013). Modelos para la prevención de bancarrotas empresariales utilizados por el sector empresarial costarricense. *Tec Empresarial*, *7*(3), 43-49. https://hdl.handle.net/2238/8187

Chudson, W. A. (1945). A survey of corporate financial structure. In *The Pattern of Corporate Financial Structure: A Cross-Section View of Manufacturing, Mining, Trade, and Construction, 1937* (pp. 1-16). NBER. https://www.nber.org/system/files/chapters/c9209/c9209.pdf

Coral & Gudiño (2013). Contabilidad universitaria. Seventh edition, McGraw-Hill. ISBN: 9789584104304.

Correa Mejía, D. A., y Lopera Castaño, M. (2019). Pronóstico de insolvencia empresarial en Colombia a través de indicadores financieros. *Panorama Económico*, *27*(2), 510–526. https://doi.org/10.32997/2463-0470-vol.27-num.2-2019-2639

Correa-Mejía, D. A., & Lopera-Castaño, M. (2020). Financial ratios as a powerful instrument to predict insolvency; a study using boosting algorithms in Colombian firms. *Estudios gerenciales*, *36*(155), 229-238. https://doi.org/10.18046/j.estger.2020.155.3588

DataScientest. 2022. “Random Forest: Bosque Aleatorio. Definición y Funcionamiento.” DataScientest. 2022. https://datascientest.com/es/random-forest-bosque-aleatorio-definicion-y-funcionamiento

Dirman, A. (2020). Financial distress: the impacts of profitability, liquidity, leverage, firm size, and free cash flow. *International Journal of Business, Economics and Law*, *22*(1), 17-25. https://ijbel.com/wp-content/uploads/2020/08/IJBEL22_205.pdf

Du Jardin, P. (2015). Bankruptcy prediction using terminal failure processes. *European Journal of Operational Research*, *242*(1), 286-303. https://doi.org/10.1016/j.ejor.2014.09.059

Duda, M., Schmidt, H., & Asgharian, H. (2010). Bankruptcy prediction: Static logit model versus discrete hazard models incorporating macroeconomic dependencies. *Lund University*, 1-60. https://www.lunduniversity.lu.se/lup/publication/1614029

Durana, P., Michalkova, L., Privara, A., Marousek, J., & Tumpach, M. (2021). Does the life cycle affect earnings management and bankruptcy?. *Oeconomia Copernicana*, *12*(2), 425-461.

Dussan Montoya, H. A., & Coy Coy, C. B. (2021). Método para la determinación de insolvencia financiera a partir de algoritmos de clasificación borrosa supervisada y no supervisada.. http://hdl.handle.net/11396/6947

Ene, A. B. (2016). Economic and Financial Analysis of a Company. *HOLISTICA Journal of Business and Public Administration*, *7*(1), 90-100.

Fahmi, Irham. "Analisis laporan keuangan (D. Handi." *ALFABETA, cv* (2020).

Faris, H., Abukhurma, R., Almanaseer, W., Saadeh, M., Mora, A. M., Castillo, P. A., & Aljarah, I. (2020). Improving financial bankruptcy prediction in a highly imbalanced class distribution using oversampling and ensemble learning: a case from the Spanish market. *Progress in Artificial Intelligence*, *9*, 31-53.

Faxas del Toro & Atucha Fuentes (2011). El análisis financiero del capital de trabajo en la empresa. *Observatorio de la Economía Latinoamericana*, (152). http://www.eumed.net/cursecon/ecolat/cu/2011

Fedorova, E., Gilenko, E., & Dovzhenko, S. (2013). Bankruptcy prediction for Russian companies: Application of combined classifiers. *Expert systems with applications*, *40*(18), 7285-7293. https://doi.org/10.1016/j.eswa.2013.07.032

Fitzpatrick, F. (1932). A Comparison of Ratios of Successful Industrial Enterprises with Those of Failed Firm. Certified Public Accountant, 6, 727-731.

Franco, Y. A. (2022). Un análisis bibliométrico de la predicción de quiebra empresarial con Machine Learning. *Odeon*, 22, 87-126. doi: https://doi.org/10.18601/17941113.n22.04

Galán-Barrera, J. A., & Torres-García, L. F. (2017). El fracaso empresarial en Colombia: Aproximación a través del modelo de Fulmer. Revista Civilizar de Empresa y Economía, 7(13). https://revistas.usergioarboleda.edu.co/index.php/ceye/article/view/1085

Garay, S.Y., Talavera, M.D., & Mendoza, W.O. (2019). Normas Internacionales de Información Financiera Analisis de la aplicación de los instrumentos financieros sección No. 11 NIIF Pymes en los estados financieros de la empresa SMAWI S.A. durante el periodo 2017. https://repositorio.unan.edu.ni/10864/

García, J. F. I., & Flores, O. (2010). Modelo probabilístico de bancarrota para bancos norteamericanos ante la recesión no reconocida del 2008. Una herramienta para la toma de decisiones. *Contribuciones a la Economía*, (2010-04). http://www.eumed.net/ce/2010a/

García-Rosales, A., & Chávez-Chávez, E. (2023). La importancia de los reportes financieros en los agronegocios como nuevas empresas. *Revista Ciencia e Innovación Agroalimentaria de la Universidad de Guanajuato*. https://doi.org/10.15174/cia.v1i1.4

García, V., Marques, A. I., & Sánchez, J. S. (2019). Exploring the synergetic effects of sample types on the performance of ensembles for credit risk and corporate bankruptcy prediction. *Information Fusion*, *47*, 88-101. https://doi.org/10.1016/j.inffus.2018.07.004

Gelashvili, V., Camacho-Miñano, M. del M., & Segovia-Vargas, M. J. (2020). Un estudio sobre el análisis económico-financiero de las empresas sociales: ¿son realmente negocios? A study of the economic and financial analysis for social firms: are they really businesses?. *Revista de Contabilidad - Spanish Accounting Review*, *23*(2), 139–147. https://doi.org/10.6018/rcsar.361531

Géron, A. (2022). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*. " O'Reilly Media, Inc.". ISBN: 9781098125974. https://www.oreilly.com/library/view/hands-on-machine-learning/9781098125967/

Gitman, L. J., Juchau, R., & Flanagan, J. (2015). Principles of managerial finance. 6th Edition. Pearson Higher Education AU. ISBN: 9781442518193. https://www.academia.edu/download/52266901/complete.pdf

Giovanni, A., Utami, D. W. and Yuzevin, T. (2020) 'Leverage dan Profitabilitas dalam Memprediksi Financial Distress Perusahaan Pertambangan Periode 2016- 2018', *Journal of*

Business and Banking, 10(1), pp. 151–167. Available at: https://journal.perbanas.ac.id/index.php/jbb/article/view/2292.

Gómez Barrera, M. A. *Ingeniería analítica para la predicción de fracaso de las microempresas en Colombia* (Master's thesis, Universidad de La Sabana). http://hdl.handle.net/10818/46535

Gupta, J., Gregoriou, A., & Ebrahimi, T. (2018). Empirical comparison of hazard models in predicting SMEs failure. *Quantitative Finance*, *18*(3), 437-466. https://doi.org/10.1080/14697688.2017.1307514

Gurnani, I., Tandian, F. S., & Anggreainy, M. S. (2021, September). Predicting Company Bankruptcy Using Random Forest Method. In *2021 2nd International Conference on Artificial Intelligence and Data Sciences (AiDAS)* (pp. 1-5). IEEE. https://doi.org/10.1109/AiDAS53897.2021.9574384

Haro Sarango, A.F., Carranza Guerrero, M.N., López Solís, O.P., Mayorga Naranjo, C.E., & Morales Ramos, K.E. (2023). Razones financieras de liquidez y actividad: herramientas para la gestión empresarial y toma de decisiones. *LATAM Revista Latinoamericana de Ciencias Sociales y Humanidades.* https://doi.org/10.56712/latam.v4i1.425

Huang, J., Wang, H., & Kochenberger, G. (2017). Distressed Chinese firm prediction with discretized data. *Management Decision*, *55*(5), 786-807. https://doi.org/10.1108/MD-08-2016-0546

Hidayat, M. A., & Meiranto, W. (2014). *Prediksi financial distress perusahaan manufaktur di Indonesia (studi empiris pada perusahaan manufaktur yang terdaftar di Bursa Efek Indonesia periode 2008-2012*) (Doctoral dissertation, Fakultas Ekonomika dan Bisnis). http://eprints.undip.ac.id/43079/

IBM. 2021. “Regresión Logística Binaria.” Regresión Logística Binaria. 2021. https://www.ibm.com/docs/es/spss-statistics/beta?topic=regression-binary-logistic.

———. n.d. “¿Qué Es Un Bosque Aleatorio?” Accessed November 28, 2023. https://www.ibm.com/es-es/topics/random-forest.

Jabeur, S. B., & Fahmi, Y. (2018). Forecasting financial distress for French firms: a comparative study. *Empirical Economics*, *54*, 1173-1186. https://doi.org/10.1007/s00181-017-1246-1

Jayasekera, R. (2018). Prediction of company failure: Past, present and promising directions for the future. International Review of Financial Analysis, 55, 196-208. https://doi.org/10.1016/j.irfa.2017.08.009

Johnpaul, C. I., Prasad, M. V., Nickolas, S., & Gangadharan, G. R. (2019). General representational automata using deep neural networks. *Data & Knowledge Engineering*, *122*, 159-180. https://doi.org/10.1016/j.datak.2019.06.004

Joshi, S., Ramesh, R., & Tahsildar, S. (2018, June). A bankruptcy prediction model using random forest. In *2018 second international conference on intelligent computing and control systems (ICICCS)* (pp. 1-6). IEEE. https://doi.org/10.1109/ICCONS.2018.8663128

Kim, M. J., & Kang, D. K. (2010). Ensemble with neural networks for bankruptcy prediction. *Expert systems with applications*, *37*(4), 3373-3379. https://doi.org/10.1016/j.eswa.2009.10.012

Lakshmi, G., Banu, K.A., Afrin, F., & Divya, C. (2021). A STUDY ON THE FINANCIAL ANALYSIS OF RELIANCE INDUSTRIES LIMITED. *International Journal of Advanced Research*. http://dx.doi.org/10.21474/IJAR01/12818

Leon, P. Y. & Vargas, J. E. (2023). *Factores que afectan el riesgo de insolvencia financiera de las empresas manufactureras de Colombia según su tamaño, mediante los Modelos Z-Score de Altman, CA Score, Springate y Fulmer durante el periodo Pre-Pandemia y Pandemia (2017-2021)*. Recuperado de: http://hdl.handle.net/20.500.12749/21244

Lev, B., Ordoñez, P. C., & Porcel, J. A. (1978). *Análisis de estados financieros: un nuevo enfoque*. Edic. Esic. Madrid, ISBN: 84-7356-016-7

Li, D., & Xia, Y. (2014). Stock Liquidity and Bankruptcy Risk. http://www.fmaconferences.org/Nashville/Papers/LiandXia_StockLiquidityandBankruptcyRisk_FMA.pdf

Liang, D., Lu, C. C., Tsai, C. F., & Shih, G. A. (2016). Financial ratios and corporate governance indicators in bankruptcy prediction: A comprehensive study. *European journal of operational research*, 252(2), 561-572. https://doi.org/10.1016/j.ejor.2016.01.012

Lin, F., Yeh, C. C., & Lee, M. Y. (2011). The use of hybrid manifold learning and support vector machines in the prediction of business failure. *Knowledge-Based Systems*, *24*(1), 95-101. https://doi.org/10.1016/j.knosys.2010.07.009

López, E. M. (2015). *Modelo de predictibilidad de quiebra en las pymes colombianas del sector comercio*. Recuperado de: http://hdl.handle.net/10726/1064

Lundqvist, D., & Strand, J. (2013). Bankruptcy prediction with financial ratios-examining differences across industries and time. https://lup.lub.lu.se/student-papers/record/3918017/file/3918043.pdf

Merwin, C. L. (1942). Financing small corporations in five manufacturing industries, 1926–36. *NBER Books*.

Moreno, E., & Bravo, F. (2018). ANALYSIS OF THE PROBABILITY OF BANKRUPTCY OF LISTED SPANISH COMPANIES. *REVISTA DE ESTUDIOS EMPRESARIALES-SEGUNDA EPOCA*, (2), 57-72. https://doi.org/10.17561/ree.v2018n2.3

Mongrut, S. M., Delgado, F. I. A., O'Shee, D. F., & Yamashiro, M. A. (2011). Determinantes de la insolvencia empresarial en el Perú. *Academia. Revista latinoamericana de administración*, (47), 126-139. https://www.redalyc.org/articulo.oa?id=71618917009

Nindita, K., & Moeljadi, N. (2014). Prediction on Financial Distress of Mining Companies Listed in BEI using Financial Variables and Non-Financial Variables. *European Journal of Business and Management*, *6*(34), 226-237. https://www.iiste.org/Journals/index.php/EJBM/article/view/17160

Ochoa Garro, Y. V., Toro Cartagena, D. C., Betancur Gallego, L. A., & Correa García, J. A. (2009). El indicador Z, una forma de evaluar el riesgo de continuidad. https://hdl.handle.net/10495/4937

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of accounting research*, 109-131. https://doi.org/10.2307/2490395

Ortiz, B. J. C. (2017). Plan de Ventas y Operaciones (PVO): estrategia para maximizar la rentabilidad de las Pymes del sector textil colombiano. https://repositorio.esumer.edu.co/handle/esumer/1898

Panigrahi, C. (2014). Relationship of Working Capital with Liquidity, Profitability and Solvency: A Case Study of ACC Limited. *ERN: Other Microeconomics: Intertemporal Firm Choice & Growth*. https://dx.doi.org/10.2139/ssrn.2398413

Parra, D. A. I., & Peluffo, D. A. B. (2022). Insolvencia de empresas colombianas a partir de indicadores financieros variables de caracterización: Aproximación a través de metodología de clasificación desbalanceada. *Económica y Financiera*, 50. https://www.supersociedades.gov.co/documents/guest/Prensa/Noticias/Publicaciones/Revistas/2022/Revista-Economica-Financiera-v1.0.pdf#page=50

Poveda, M. U. (2019). Riesgo de crédito: Evidencia en el sistema bancario ecuatoriano. *Bolentín de Coyuntura*, (23), 4-9. https://doi.org/10.31243/bcoyu.23.2019.842

Prats, G. M., Ramos, M. A., Rodríguez, W. B. L., & Gutiérrez, G. E. M. (2022). IMPORTANCIA DE LAS AUDITORIA EXTERNAS PARA EL ANÁLISIS EFECTIVO DE LOS ESTADOS FINANCIEROS. *Revista de Investigación Académica sin Frontera*, (37), 10. https://dialnet.unirioja.es/servlet/articulo?codigo=8450938

Ramírez Alcócer, D. I. (2014). Modelo de administración del capital de trabajo, para maximizar los resultados financieros y mejorar la liquidez de las PYMES. http://repositorio.puce.edu.ec/handle/22000/11729

Rahim, A.H. (2021). Evaluating the Financial and Administrative Performance using Financial Analysis Methods Field Study in a Selected Iraqi Company. *International Journal of Science and Management Studies (IJSMS).* https://www.ijsmsjournal.org/2021/volume-4%20issue-6/ijsms-v4i6p113.pdf

Romero Espinosa, Fredy. (2013). Alcances y limitaciones de los modelos de capacidad predictiva en el análisis del fracaso empresarial. *AD-minister*, (23), 45-70. Retrieved December 21, 2023, from http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1692-02792013000200004&lng=en&tlng=es

Ruiz, H. A. (2015). *Modelo de predicción de punto de quiebra de las empresas manufactureras Pymes en Colombia*. Recuperado de: http://hdl.handle.net/10726/1086.

Sandoval Serrano, L. J. (2018). Algoritmos de aprendizaje automático para análisis y predicción de datos. *Revista Tecnológica; no. 11*. http://www.redicces.org.sv/jspui/bitstream/10972/3626/1/Art6_RT2018.pdf

Septiani, Ni M. I., and I. M. Dana. "Pengaruh Likuiditas, Leverage, dan Kepemilikan Institusional terhadap Financial Distress pada Perusahaan Property dan Real Estate." *E-Jurnal Manajemen Universitas Udayana*, vol. 8, no. 5, 2019, doi:10.24843/EJMUNUD.2019.v08.i05.p19

Sepúlveda, S. P. (2021). DE LA CONSOLIDACIÓN DE LA TERCERIZACIÓN A LA PROBLEMÁTICA DEFINICIÓN DE EMPRESA EN CHILE. *Caderno CRH*, *34*, e021034. https://doi.org/10.9771/ccrh.v34i0.45096

Sermpinis, G., Tsoukas, S., & Zhang, Y. (2023). Modelling failure rates with machine-learning models: Evidence from a panel of UK firms. *European Financial Management*, *29*(3), 734-763. https://doi.org/10.1111/eufm.12369

Shahdadi, K. M., Rostamy, A. A., Sadeghi Sharif, S. J., & Ranjbar, M. H. (2020). Intellectual capital, liquidity, and bankruptcy likelihood. *Journal of Corporate Accounting & Finance*, *31*(4), 21-32. https://doi.org/10.1002/jcaf.22460

Shin, K. S., Lee, T. S., & Kim, H. J. (2005). An application of support vector machines in bankruptcy prediction model. *Expert systems with applications*, *28*(1), 127-135. https://doi.org/10.1016/j.eswa.2004.08.009

Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *The journal of business*, *74*(1), 101-124. https://doi.org/10.1086/209665

Simbaña Gualavisí, B. E. (2020). *La relación entre riesgo de liquidez y riesgo de crédito en una entidad de microfinanzas del Ecuador* (Bachelor's thesis, Quito: EPN, 2020.). http://bibdigital.epn.edu.ec/handle/15000/20799

Sinelnikova-Muryleva, E. V., Gorshkova, T. G., & Makeeva, N. V. (2018). Default forecasting in the Russian banking sector. *Economic Policy*, *13*(2), 8-27. https://repec.ranepa.ru/rnp/ecopol/ep1811.pdf

Smith, R. F., & Winakor, A. H. (1930). A test analysis of unsuccessful industrial companies. *Bureau of Business Research, Bulletin*, *31*. University of Illinois.

Smith, R. F., & Winakor, A. H. (1935). Changes in the financial structure of unsuccessful industrial corporations. *Bureau of Business Research*, Bulletin No. 51. Urbana: University of Illinois Press.

Somoza, A. L., & Vallerudú, J. C. (2009). Una comparación de la selección de los ratios contables en los modelos de predicción de la insolvencia empresarial. *Economía industrial*, (373), 153-168. https://dialnet.unirioja.es/servlet/articulo?codigo=3108051

Solórzano-Quito, D. E., & Vásconez-Acuña, L. G. (2021). Estrategias financieras y contables para el fortalecimiento de la liquidez en la Corporación Agroempresarial Coagro. *Cienciamatria*, *7*(2), 508-537. https://doi.org/10.35381/cm.v7i2.520

Sucipto, A. W., & Muazaroh, M. (2017). Kinerja rasio keuangan untuk memprediksi kondisi financial distress pada perusahaan jasa di Bursa Efek Indonesia periode 2009-2014. *Journal of Business & Banking*, *6*(1), 81-98. http://dx.doi.org/10.14414/jbb.v6i1.893

Sunarjanto, N. A., Roida, H. Y., & Widyaningdyah, A. U. (2016). Predicting Business Failure for Small Medium Enterprises (SMEs): The Role of Financial Ratios. *The International Journal of Business & Management.*, *4*(11), 64-69. http://repository.ukwms.ac.id/id/eprint/14173

Talavera, M.D., Rivera, F.A., & Mejía, A.L. (2016). Importancia de la aplicación de indicadores financieros en la toma de decisiones de la empresa New Cigars S.A. durante el periodo 2013-2014. https://repositorio.unan.edu.ni/1863/

Terreno, D. D., Peréz, J. O., & Sattler, S. A. (2020). La relación entre liquidez, rentabilidad y solvencia: Una investigación empírica por el modelo de ecuaciones estructurales. *Contaduría Universidad de Antioquia*, (77), 13-35. https://doi.org/10.17533/udea.rc.n77a01

Thakor, A.V. (2018). Post-Crisis Regulatory Reform in Banking: Address Insolvency Risk, Not Illiquidity! *Monetary Economics: Central Banks - Policies & Impacts eJournal*. https://doi.org/10.1016/j.jfs.2018.03.009

Tinoco, M. H., & Wilson, N. (2013). Financial distress and bankruptcy prediction among listed companies using accounting, market and macroeconomic variables. *International review of financial analysis*, *30*, 394-419. https://doi.org/10.1016/j.irfa.2013.02.013

Umaña, C. E. (2022). DETERMINANTES DE LOS FLUJOS DE EXPORTACIONES PARA LAS EMPRESAS EN COSTA RICA: UN ANÁLISIS DE SUPERVIVENCIA. *Revista De Ciencias Económicas*, *30*(2). https://doi.org/10.15517/rce.v30i2.8035

Valencia, C., Cabrales, S., Garcia, L., Ramirez, J., & Calderona, D. (2019). Generalized additive model with embedded variable selection for bankruptcy prediction: Prediction versus interpretation. *Cogent Economics & Finance*, *7*(1), 1597956.

Valladares, E. V., & Flores, J. L. D. (2005). Análisis de razones financieras en la empresa lechera intensiva: un estudio de caso en el altiplano mexicano. *Veterinaria México*, *36*(1), 25-40. http://www.redalyc.org/articulo.oa?id=42336103

Viana, A. D. & 1140861895 (2021). *Predicción de quiebra en hospitales Colombianos: modelo de gestión de riesgo financiero.* Recuperado de: http://hdl.handle.net/10726/4393

Vinet, L., & Zhedanov, A. (2011). A 'missing' family of classical orthogonal polynomials. *Journal of Physics A: Mathematical and Theoretical*, *44*(8), 085201. https://doi.org/10.1088/1751-8113/44/8/085201

Virgilio, G. P., Caro, N. L. M., Sarmiento, R. N. M., Rivera, J. D., & Díaz, Í. R. (2022). Credit risk and profitability of short-term deposit at Savings and Credit Cooperatives. The case of Peru. *REVESCO: revista de estudios cooperativos*, (142), 11. https://dx.doi.org/10.5209/REVE.84396

Viswanathan, P. K., Srinivasan, S., & Hariharan, N. (2020). Predicting financial health of banks for investor guidance using machine learning algorithms. *Journal of Emerging Market Finance*, *19*(2), 226-261. https://doi.org/10.1177/0972652720913478

Wang, G., & Ma, J. (2012). A hybrid ensemble approach for enterprise credit risk assessment based on Support Vector Machine. *Expert Systems with Applications*, *39*(5), 5325-5331. https://doi.org/10.1016/j.eswa.2011.11.003

Wang, G., Ma, J., & Yang, S. (2014). An improved boosting based on feature selection for corporate bankruptcy prediction. *Expert Systems with Applications*, *41*(5), 2353-2361. https://doi.org/10.1016/j.eswa.2013.09.033

Wilson, R. L., & Sharda, R. (1994). Bankruptcy prediction using neural networks. *Decision support systems*, *11*(5), 545-557. https://doi.org/10.1016/0167-9236(94)90024-8

Wulandari, A. F. (2019). *Pengaruh Profitabilitas, Leverage, Ukuran Perusahaan, dan Umur Perusahaan Terhadap Ketepatan Waktu Penyampaian Laporan Keuangan (Studi Empiris Pada Perusahaan Manufaktur yang Terdaftar di Bursa Efek Indonesia Periode 2015-2018)* (Doctoral dissertation, Skripsi, Universitas Muhammadiyah Magelang). http://eprintslib.ummgl.ac.id/676/

Xu, X. (2019). Forecasting bankruptcy with incomplete information. *RMI Working Paper*, No. 55024, pp. (1) – (42). https://mpra.ub.uni-muenchen.de/id/eprint/55024

Yousaf, U. B., Jebran, K., & Wang, M. (2022). A Comparison of Static, Dynamic and Machine Learning Models in Predicting the Financial Distress of Chinese Firms. *Romanian Journal of Economic Forecasting*, *25*(1), 122. https://ssrn.com/abstract=3972280

Zacharakis, A. L., Meyer, G. D., & DeCastro, J. (1999). Differing perceptions of new venture failure: a matched exploratory study of venture capitalists and entrepreneurs. *Journal of small business management*, *37*(3), 1-14. https://www.proquest.com/openview/2a5468adabf749154bc46a47e6deb034/1?pq--origsite==gscholar&cbl==49244

Zeng, J. (2013). Financial Analysis of the Company Metro AG. Online, Bakalářská práce. Ostrava: Vysoká škola báňská - Technická univerzita Ostrava. http://hdl.handle.net/10084/97649

Zięba, M., Tomczak, S. K., & Tomczak, J. M. (2016). Ensemble boosted trees with synthetic features generation in application to bankruptcy prediction. *Expert systems with applications*, *58*, 93-101. https://doi.org/10.1016/j.eswa.2016.04.001